Defense Technical Information Center Compilation Part Notice

ADP011354

TITLE: Some Contributions to Wavelet Based Image Coding

DISTRIBUTION: Approved for public release, distribution unlimited

This paper is part of the following report:

TITLE: Input/Output and Imaging Technologies II. Taipei, Taiwan, 26-27 July 2000

To order the complete compilation report, use: ADA398459

The component part is provided here to allow users access to individually authored sections of proceedings, annals, symposia, etc. However, the component should be considered within the context of the overall compilation report and not as a stand-alone technical report.

The following component part numbers comprise the compilation report:
ADP011333 thru ADP011362

Some contributions to wavelet based image coding

Yi-Wei Li, Kuo-Shu Chang, Loon Shan Yan and Day-Fann Shen
Department of Electrical Engineering
Yunlin University of Science and Technology
Taiwan

ABSTRACT

Four key issues in wavelet zero-tree based image coding are investigates and presented, they are (1) Fast wavelet transform that save 1/2 and 3/4 processing for one dimensional signal and two dimensional signals respectively. (2) The selection of the best wavelet filters that yields best performance (PSNR vs. Bit rate) for most common seen images. (3) Recommendation of number of wavelet scales (or frequencies) for image coding by experiments and analysis.

Keywords: wavelet transform, image coding, wavelet scales, filters.

1. INTRODUCTION

Wavelet zero-tree based approach has been adopted in JPEG2000 image coding standard [1]. In this paper, We discuss four important issues in a wavelet zero-tree based image coding and present our solutions for improvements. Firstly, Wavelet decomposition/synthesis are the basic process in any wavelet based image coding, it is highly desired to speed up the process without sacrifice the precision. By integrating the convolution and the sub-sampling process, our approach is able to speed up the wavelet transform up to 1/2 for one dimensional signal and 3/4 for two dimensional signal. The correctness of the results is verified using MATLAB wavelet toolbox.

Secondly, it is known that the main reason that discrete wavelet transform (DWT) outperforms discrete cosine trasnform (DCT) in image coding is that wavelets of finite duration forms a set of better basis than the periodic cosin basis in representing the signals and the images. Many wavelet filters have been proposed [2] [3], it is interested to know which wavelet filters are the best for image coding. We have tested all 77 sets of wavelet coefficients provided in MATLAB wavelet toolbox (version 1) for their performances (Bit-rates vs. PSNRs) using a zero-tree based coding technique [4]. The results indicate that two set of wavelet filters Bior-4.4 and Bior-5.5 are consistently the best for three test images of different complexity: Lena, Goldhill and Pepper. Both ing technique [4]. The results indicate that two sets of wavelet filters are bi-orthogonal and linear phased filters. It is known that Bior-4.4 filter is also recommended for the coding of finger print by CIA [5].

Finally, one wavelet decomposition generate four sub-bands (LL, HL, LH, HH) of the same scale, the octave wavelet subbands of different scales can be obtained by decomposing the lowermost subband (LL) repeatedly, an wavelet subbands of three scales are shown in Figure 1 . One key parameter in wavelet transform is the optimal number of scales for image coding. We investigate this issue via performance evaluation with a brief analysis. We recommend that for image of higher complexity, like Goldhill, 3 scales are the best while for images of medium and low complexity (Lena and Pepper), 4 scales of decomposition is the best choice.

2. FAST WAVELET ANALYSIS/SYNTHESIS

Signal analysis and synthesis based on the wavelet can be efficiently implemented using a pair of QMF (Quadratic Mirror Filters) filters as proposed by Mallat[2]. A one dimensional one scale (level) wavelet analysis and synthesis process is shown in Figure 2. Wavelet subbands with octave scales can be obtained by successively decomposing the lowest frequency subband. Since the bandwidth of the low-passed and high-passed signal is halved, they can be down-sampled by a factor of 2. Let m be the length of the input signal X(n) and n be the length of an analysis filter, m is usually an even number while n can be either even or odd. The length of the signal immediately after circular convolution (X(n) is regarded as an periodic signal) with the analysis filter and before down-sampling is $m+n-1$. It is normally required that length of $Y_l(n)$ and $Y_h(n)$ be $m/2$, therefore, the filter output must be truncated to

In *Input/Output and Imaging Technologies II*, Yung-Sheng Liu, Thomas S. Huang, Editors, Proceedings of SPIE Vol. 4080 (2000) • 0277-786X/00/$15.00

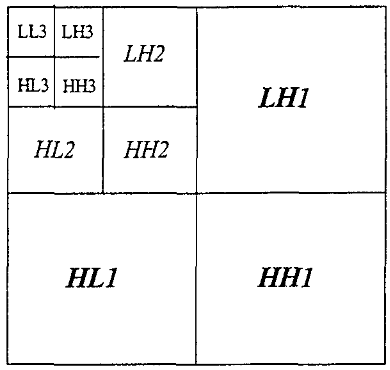

Figure 1. 3-scale wavelet decomposition

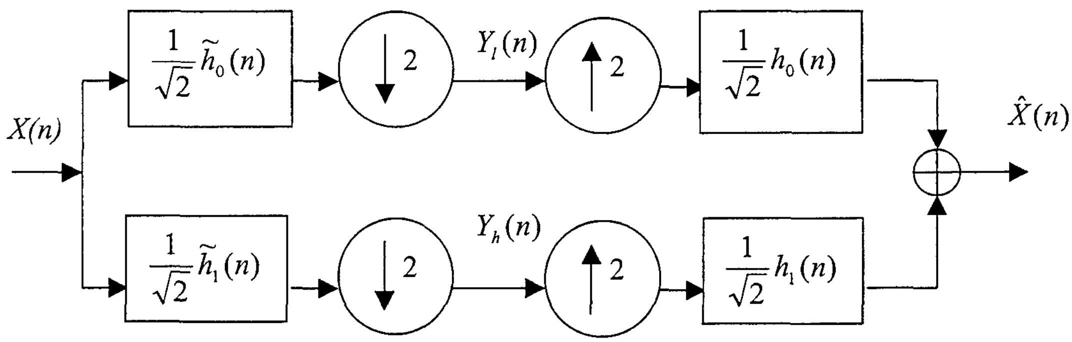

Figure 2. One dimensional one scale (level) wavelet analysis and synthesis process using QMF.

m before down-sampling. A common practice to achieve this requirement is by removing (n-1)/2 samples at both end of the filtered output if n is even or remove $n/2$ samples at one end and $n/2$-1 at the other end. The length truncated signal is then down-sampled to obtain $Y_l(n)$ and $Y_h(n)$. The function "dwt2per" in MATLAB wavelet toolbox is designed to perform the truncation.

We proposed a method to improve the process of generating $Y_l(n)$ and $Y_h(n)$, the idea is that convolution and down-sampling can be done in one process. Figure 3 is an example used for illustration, here n=4 is used. Note that if the length of the analysis and synthesis filters are not equal, zeros be added to the shorter filter. X(n)= [X(1)....X(m)] is treated as a periodic signal of period m, f(n) = [f(1)...f(4)] is the reversed ordered filter coefficients for the convolution process.

The low-pass or high-pass filter output Y(k) = $\sum_{j=1}^{j=n} f(j)x(2(k-1)+j)$ for k=1..$m/2$. Note that only $m/2$ points of outputs are calculated. This simple idea reduces the computations to $(1/2)^k$ of the original method. Moreover, the signal truncating problem is now eliminated. The reconstruction process includes up-sampling (inserting zeros between signal data) to form a periodic signal Y'(n) of period m. To reconstruct the signal X'(k), X'(k)= $\sum_{j=1}^{Round[n/2]} f(2j-1)Y'(k+2j-2)$ if k is odd and X'(k)= $\sum_{j=1}^{Round[n/2]} f(2j-1)Y'(k+2j-1)$, if k is even. The example with filter length n=4 is shown in Figure 4,where f(1)-f(4) are the reverse ordered reconstruction filter coefficients. The output signals from low-pass and high-pass reconstruction filters are then summed up to reconstruct the original signal.

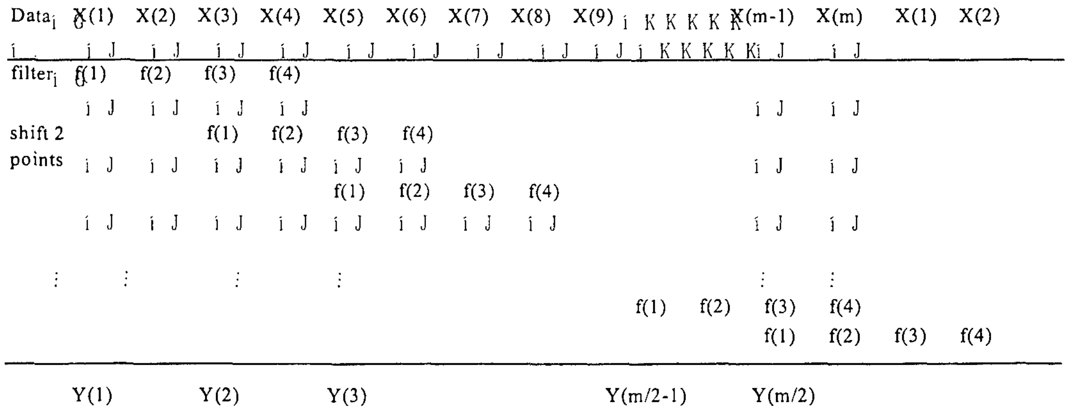

Figure 3: Example of integrating circular convolution and down-sampling in the analysis process

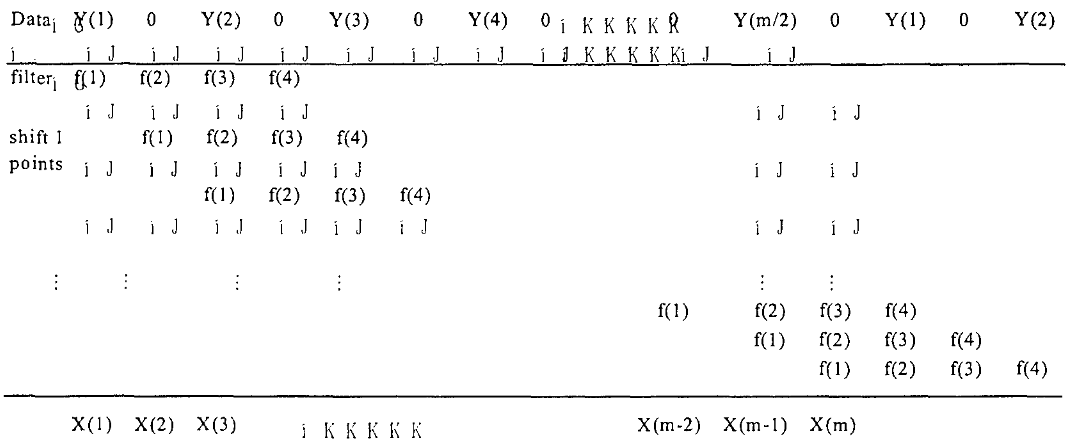

Figure 4: Example of the synthesis process (corresponding to the analysis process in Figure 3)

3. SELECTION OF WAVELET FILTERS FOR ZERO-TREE BASED IMAGE CODING

There are many wavelets proposed, in fact, there are 77 sets of wavelet filters collected in MATLAB wavelet toolbox. Each wavelet filter has several inherited properties that may affect the coding gain, reconstructed image quality and computational complexity. These properties are (1) separability, (2) linear phase, (3) Orthogonality, (4) regularity and (5) filter size (length).

Separable wavelet filter and separable down-sampling allow very efficient implementation of wavelet transform, since the 2-D filtering can be break down to two cascaded 1-D filtering process, the total computations is reduced from $O(N^4)$ to $O(N^3)$. However, the drawback of separable wavelet transform is that only rectangle pieces of the spectrum can be isolated, this is because separable 2-D filtering is the product of 2 1-D filters. Different shapes of spectrum other than rectangle would require non-separable 2-D filters, which allow better coding performance at the cost of higher computational complexities and may have the stability problem. [6]. In practice, most wavelet transforms are using separable filters, all the 77 sets of wavelet filters in MATLAB wavelet toolbox are all separable filters.

Phase distortion is more visible than magnitude distortion in a reconstructed image. Without linear phase distortion around edges is very visible. As a result, it is strongly advocated that linear phase filter be used in the sub-band coding.

Orthogonal wavelet filters implement "unitary transform" between input and the decomposed sub-bands. It implies that energy, distortion as well as bit rate are conserved between the input signal and the decomposed sub-band signals. The bit allocation algorithm can be implemented easily using this property. However, it has proven that linear phase and orthogonalilty are mutual exclusive in a separable FIR system. Linear phase using orthogonal filter can be achieved under the following conditions. (1) An orthogonal filter of sufficient length can be made almost linear phase. (2) Non-separable orthogonal filters may have linear phase and (3) Orthogonal IIR filter allow linear phase. In practice, most implementation of wavelet transforms adopt well designed Bi-orthogonal filter for linear phase while keep the conservation property as close to an orthogonal filter as possible.

An orthogonal filter with a certain number of zeros at the aliasing frequencies (π in two channels case) is called regular if its iteration tends to a continuous function. A filter of high regularity improves coding gain the compression artifacts is less objectional. Filters of low regularity cause poor coding performance, moderate regularity improves the performance significantly, however, higher regularity can only improve the performance a little. For bi-orthogonal filtering, only either the analysis or the synthesis filter can be regular. To minimize the visibility of objectional artifacts, it is preferred to have a regular synthesis filter.

Wavelet filter of higher regularity requires longer filters, however, longer filters have the following drawbacks, (1) It requires more computations. (2) It tends to spread coding error around (3) Longer filter tends to have more zero-crossings which causes more ringing artifacts around edges. For image compression, shorter and smoother filters are preferred [6].

To select the best wavelet filter available for zero-tree based image coding, we conduct an experiment to evaluate the performance for all 77 sets of wavelet filters collected in MATLAB wavelet toolbox version 1. The 256x256 Lena image and JZW (JND based zero-tree wavelet) as proposed by Shen and Yan in 1998 [4] are used in the initial performance evaluation process. Among the best 9 sets of wavelet filters (in PSNR), 6 sets are biorthogonal filters (Bior2.2, bior2.4, Bior2.6, bior4.4, bior5.5, bior6.8) and 3 sets are symlet (sym4, sym5, sym6). The PSNR vs. bit rate curves for these 9 set filters are shown in Figure 5.

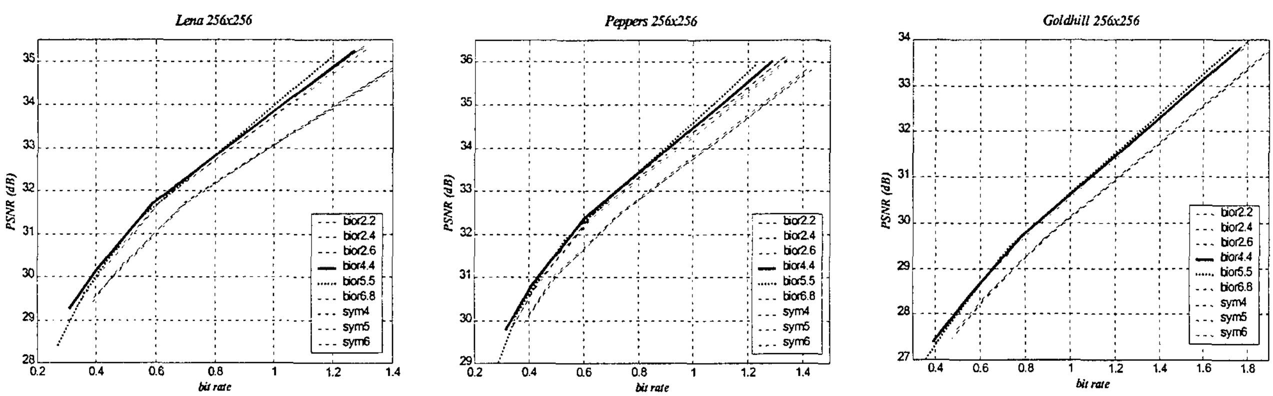

Figure 5. Performance (PSNR vs. Bit-rate) of the best 9 wavelet filters in MATLAB wavelet toolbox

Among the 9 sets of the best wavelet filters, bior4.4 and bior5.5 consistently outperform the others for Lena, Pepper and Goldhill images in the range of PSNR 28 to 35. By further examining the detailed performance data for bior4.4 and bior5.5 in Table I, we found that bior4.4 performs better in the range of low to medium bit rate(below 0.8 bpp) while 5.5 performance better for higher bit rate (0.8 bpp and above). The filter coefficients for bior4.4 (9/7 filter) and bior5.5 (11/9 filter) are listed in Table II and Table III.

Table I. Bit rate vs. PSNR data for Bior4.4 and Bior5.5 filters.

factor \ Bit rate PSNR \ Complexity		Goldhill (high)		Lena (median)		Peppers (Low)	
		bior4.4	bior5.5	bior4.4	bior5.5	bior4.4	Bior5.5
£ p1	bpp	1.7635	1.7341	1.2725	1.2046	1.2898	1.2270
	PSNR(dB)	33.8227	33.8407	35.2504	35.1433	36.0072	35.9261
£ p3	bpp	0.7805	0.7253	0.5864	0.5365	0.6025	0.5296
	PSNR(dB)	29.7120	29.4590	31.7021	31.2900	32.3795	31.8880
£ p5	bpp	0.5274	0.4650	0.4016	0.3548	0.4096	0.3568
	PSNR(dB)	28.2795	27.8079	30.1586	29.5543	30.8100	30.1360
£ p7	bpp	0.3881	0.3331	0.3092	0.2659	0.3147	0.2737
	PSNR(dB)	27.3825	26.8423	29.2472	28.3659	29.7608	28.8183

Table II. Filter coefficients for Bior4.4 (9/7 taps QMF)

n		0	±1	±2	±3	±4
Decomposition Filter	$\tilde{h}_0(n)$	-0.557543	0.295636	0.02877	-0.045636	0
	$\tilde{h}_1(n)$	0.602949	0.266864	-0.07823	-0.016864	0.026749
Reconstruction Filter	$h_0(n)$	-0.602949	0.266864	0.78223	-0.016864	-0.026749
	$h_1(n)$	0.557543	0.295636	-0.02877	-0.045636	0

Table III. Filter coefficients for Bior5.5 (11/9 taps QMF)

n		0	±1	±2	±3	±4	±5
Decomposition Filter	$\tilde{h}_0(n)$	0.6360	-0.3372	-0.0661	0.0967	-0.0019	-0.0095
	$\tilde{h}_1(n)$	0.5209	0.2444	-0.0385	0.0056	0.0281	0
Reconstruction Filter	$h_0(n)$	0.5209	-0.2444	-0.0381	-0.0056	0.0281	0
	$h_1(n)$	0.6360	0.3372	-0.0661	-0.0967	-0.0019	0.0095

The selection of Bior4.4 and Bior5.5 filters are consistent with the above requirements analysis of an idea wavelet filters in image coding. Both are separable, bi-orthogonal, linear phase, short and smooth with moderate regularity.

4. NUMBER OF WAVELET SCALES IN WAVELET ZERO-TREE BASED IMAGE CODING

Theoretically, octave subbands can be obtained through repeated wavelet analysis on the low-frequency subband until a single pixel is reached [6], thus a maximum number of scales for a MxM image is $N_{Max\_scales} = \mathrm{Log}_2 M$. It is nature to ask the question: How many wavelet scales is most suitable for image coding ? There are few papers discuss this problem, SPIHT [7] arbitrarily chooses 5 scales, while Rajala etal. [8] choose 3 scales. In this paper, we investigate the relation between number of wavelet scales and coding performance in image coding. A zero-tree based coding technique JZW [4] is adopted for the performance evaluation purpose. The wavelet transformed coefficients in each sub-band are quantized by a JND weighted step size $\Delta_{subband}$, this process is called JND_SQ (or JND based Scalar Quantization). $\Delta_{subband}$ for each sub-band are derived from extensive experiments and has larger value for higher frequency sub-band and has smaller value for lower frequency sub-band. Thus, wavelet coefficients in higher frequency (smaller scale) subbands are quantized more coarsely while the lower

frequency (larger scale) subbands are quantized relatively finer. After JND_SQ, wavelet coefficients smaller than 1/2 of the step size are quantized to zeros. As a result, many zeros are produced in each sub-band, especially at higher frequency sub-bands where step size is large and coefficients values are smaller. It is known that the more zeros, the higher coding efficiency in zero-tree based image coding. In addition, it is important to know that the JND step sizes are carefully derived that the reconstructed image after JND_SQ maintaining a visually loss-less quality even under viewing condition in dark room and at any viewing distance. Although JZW is designed to optimize the perceptual image quality, several zero-tree based image compression techniques are proposed to enhance the coding performance, as a result, JZW outperforms SPIHT even in terms of MSE or PSNR. However, it is noted that JZW does not implement the embedded property as in EZW and SPIHT and that the embedded property can be achieved by passing the JND quantized wavelet coefficients to EZW or SPIHT.

It is noted that the lowest frequency (the coarsest scale) band (LL band) is the most important subband to human perception. Also, most coefficients in LL band have very large values and unlikely to be zeros, it is not efficient to include LL band in the zero-tree scanning. For these two reasons, JZW encodes the LL band separately using loss-less DPCM. Wavelet coefficients in other higher frequency subbands can be efficiently encoded using our zero-tree encoding scheme which is derived from EZW and [improved version of EZW by SAID and PERALMAN, as indicated by song in 4/28/2000 presentation]. In JZW, each zero-tree is encoded by list of coefficient states (LCS) and list of coefficient values (LCV). To simplify the implementation, JZW omits the sophisticated embedded property, i.e. JZW encodes the full value (JND quantized) of the coefficients in one pass rather than the bit plans in multiple passes. JZW also simplifies the 4 possible states to 3 possible states for zero-trees. The three states are {ST (Significant Tree) , SR (Significant Root) and ZTR (ZeroTree Root) }. A ST coefficient has at least one non-zero descendent; a SR coefficient is non-zero itself but all descendents are zeros; A ZTR coefficient and all its descendents are zeros. The children of a ST coefficient may have three possible states {ST, SR and ZTR}; Children of {SR and ZTR} must be ZTRs and their descendents must be all zeros and can be skipped in the coding process. It is noted that the more zero coefficients the higher coding efficiency and that JND_SQ can effectively reduce those coefficients smaller than 1/2 $\Delta_{subband}$ into zeros without degrading the perceptual quality.

Since we are interested in determining the wavelet scales, the wavelet decomposition of 3, 4 and 5scales (corresponding to total of 10, 13,16 wavelet subands) are generated for a set of 6 test images (Goldhill, Lena, Pepper of 512x512 and 256x256). The performance (Bit rate vs. image quality) of JZW on each test image of different scales are recorded. Table IV lists the best wavelet scales for the 6 test images.

Table IV. Best wavelet scales for the 6 test images

complexity / size / Scale	Goldhill (high)	Lena (moderate)	Peppers (Low)
512x512	3	4	4
256x256	3	3	3

Based on the data in Table IV, we conclude that 3 or 4 scales are most suitable for common seen images. For images of lower resolution (256x256) and images of higher complexity (Goldhill), 3 scales (10 subbands) is the best; For images of higher resolution (512x512) and moderate or lower complexity, 4 scales (13 subbands) is the best.

We further investigate the reasons behind the optimal decomposition levels. Consider a 5-levels wavelet zerotree shown in

Table V. If a node at level 5 (the lowest frequency band) is an SR(condition 1), then 4-level decompositions (4 scales) for the same image would require one more symbol than 5 level decomposition. However, if a node at level 5 is ZTR or SR, then 4-level decomposition (4 scales) can save at least three symbols than 5-level decomposition (5 scales). There are tradeoffs in terms of required number of symbols. How to choose the optimal levels of wavelet decomposition depends on the percentage of condition 1 and condition 2 nodes.

After N level decompositions, if the percentage of condition 2 nodes (ZTR or SR) and condition 1 (ST nodes) are 25% and 75% respectively, then N level (Nscales) and N-1 level decomposition requires about the same number of symbols. Therefore, the rule of thumb is that *if the percentage of condition 2 (ZTR or SR) is less than 25% after N levels decomposition, then the optimal levels of decomposition is **N-1**.* To illustrate, we take 512x512 and 256x256 "Lena" as an example. Table VII shows the percentage of condition 2 for different levels of wavelet decomposition. We may find the optimal levels of wavelet decomposition for 512x512 as well as 256x256 Lena using the above rule of thumb. ST node often appears on the edges while SR/ZTR appears in smooth areas of an image. The more complex an image is, the more edges it has. So there is fewer percentage of ZTR /SR in a complex image, therefore, fewer wavelet decomposition levels are necessary.

In general, most of the commonly seen natural images have low or medium complexity. Therefore, 3-levels wavelet decomposition is recommended for 256x256 or the similar sizes such as CIF (352x288) or QCIF (176x144), while 4-levels wavelet decomposition is recommended for 512x512 images or the similar sizes respectively.

Table V. Example of required symbols

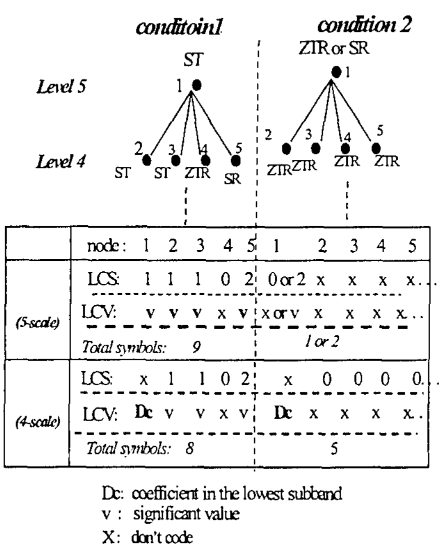

Table VI. Percentage of condition 2

Lena

	512x512	256x256
Level 1	99.83 %	99.26%
Level 2	92.55%	85.74%
Level 3	69.45%	52.96%
Level 4	36.16%	18.16%
Level 5	8.07%	1.56%

5. CONCLUSION

We have presented our contributions on four key issues in wavelet zero-tree based image coding. They are (1) Fast wavelet transform that save 1/2 and 3/4 processing for one dimensional signal and two dimensional signals respectively. (2) The selection of the best wavelet filters that yields best performance (PSNR vs. Bit rate) for most common seen images. (3) Recommendation of number of wavelet scales for image coding by experiments and analysis

6. REFERENCES

[1] www.JPEG.org//JPEG2000.htm
[2] Stephane G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation, " *IEEE Trans. Pattern Analysis and Machine Int elligence*, vol. 11, no.7,pp.674-693,July 1989.
[3] Ingred Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Tran. Information*

Theory, vol. 36, no. 5,Sep.1990.
[4] Day-Fann Shen and Loon-shan Yan, "JND measurements and wavelet based image coding", *SPIE* International Optoelectronics Exposition, July 1998
[5] A. B. Watson and G. Y. Yang, "Visibility of Wavelet Quantization Noise," *IEEE Trans.* Vol. 6, no. 8, August 1997
[6] Martin Vetterli and Jelena Kovacevic, "Wavelets and Subband Coding"
[7] Amir Said and William, "A new, Fast, and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees". *IEEE Transactions on circuits and systems for video technology,* vol. 6, pp. 243-250, June 1996.
[8] Robert E. Van Dyck and Sarah A. Rajala, "Subband/VQ Coding of Color Images With Perceptually Optimal Bit Allocation". *IEEE Transactions on circuits and systems for video technology,* vol. 4, pp. 68-82 , Feb. 1994.